用户搜寻意图和搜寻策略选择的关联机制*

■ 袁红
南通大学经济与管理学院 南通 226019

摘要：［目的/意义］搜寻意图是信息搜寻过程的起点，搜寻策略是搜寻行为的规划，两者的关联机制是搜寻行为动态发展规律的集中体现，研究搜寻意图和搜寻策略选择的关联机制有利于深入了解搜寻行为产生及变化的原因和过程。［方法/过程］通过问卷收集搜寻意图与搜寻策略数据，使用 UCINET 研究不同意图下的策略选择偏好，使用 SPSS modeler 探寻不同意图下的策略组合，采用 ROST CM 对基于不同意图的信息偶遇现象进行研究。［结果/结论］搜索和浏览是用户信息搜寻时最常使用的搜寻策略；不同的搜寻意图对应的用户搜寻策略选择偏好呈现明显差异，也表现出不同的搜寻策略组合模式；信息偶遇比较大概率地发生在信息查找、资源下载和建议咨询情境中。

关键词：搜寻意图 搜寻策略 关联机制
分类号：G252
DOI：10.13266/j.issn.0252-3116.2019.22.006

1 引言

用户导向的信息搜寻行为研究旨在了解人们发现和获取不同信息资源的方式，并试图探究行为规律背后的复杂机制。信息搜寻行为始于搜寻意图，搜寻意图包含信息需求、搜寻目标和搜寻动机等；用户对搜寻行为的规划即搜寻策略，搜寻策略居于行为要素的核心位置，两者均是信息搜寻行为研究的重要内容。

搜寻意图和搜寻策略选择的关联机制是用户和资源之间的交互，它是搜寻意图的外在表现，是搜寻策略的内在动因，是搜寻行为的动态发展和实现过程。长期以来，相关研究倾向于将信息搜索过程自身割裂开来，着眼于满足一个独立的、具体的信息需求。本研究以用户的搜寻意图和搜寻策略之间的联系为切入点，研究搜寻过程中行为的变化，揭示网络环境下信息需求和搜寻行为的特征和规律，具有重要的理论意义。

从用户信息搜寻实践来看，随着搜索引擎和社会化媒体的广泛应用，加之智能手机、平板等移动设备逐渐取代桌面电脑，可穿戴设备具备了信息交流功能，智能助手（如 Siri）能够主动提供用户需要或感兴趣的信息，人类信息行为正在变得无比丰富且难以预测。真实的搜寻意图和策略选择是用户行为画像的重要内容。此外，在信息爆炸的今天，尽管更多的信息以及更多的信息访问途径将信息的可获得性提升到前所未有的水平，但用户仍然面临着非常严峻的信息过剩和无用信息问题，深入研究用户搜寻行为的习惯和意向是提升用户搜寻效率和体验的关键，能够为知识服务系统和搜寻系统的建立和完善提供支撑，具有重要的应用价值。

2 相关研究

2.1 用户搜寻意图研究

学界对搜寻意图的研究呈方兴未艾之势，近年来SIGIR、WWW 等重要国际会议日益重视搜寻意图分类及意图识别研究。搜寻意图的分类研究首先围绕搜寻目标开展。A. Broder[1] 从用户需求出发，划分了信息类（informational）、事务类（transactional）以及导航类（navigational）搜寻意图，该分类具有广泛的接受度和良好的适用性。同时，一些学者出于细粒度研究的需要，尝试从目标性以外的维度细化搜寻意图。V. B. Nguyen 等[2] 从模糊性、权威敏感度、时间和空间敏感度 4 个维度定义搜寻意图。但是，从具体研究问题出发，在基本分类的基础上构建科学完备的搜索意图分类体系仍将是持续的热点。目前搜寻意图识别研究主要是利用检索式、搜寻结果以及搜寻行为数据来提取

* 本文系国家社会科学基金一般项目"用户探索式搜索行为的意图 - 策略关联模式研究"（项目编号：17BTQ087）研究成果之一。
作者简介：袁红（ORCID：0000-0003-2865-640X），教授，博士，硕士生导师，E-mail：yh@ntu.edu.cn。
收稿日期：2019-01-22 修回日期：2019-08-31 本文起止页码：49-57 本文责任编辑：易飞

和分析搜寻行为特征,N. J. Belkin 等[3]统计了查询词和查询长度;D. Vallet 等[4]对搜寻结果中的实体类别加以排序;张晓娟[5]利用日志分析了用户查询重构行为,但如何将识别出的意图应用在搜索引擎优化中却没有引起足够的关注。

2.2 信息搜寻策略研究

信息搜寻策略研究经历了 Moves→Tactics→ISSs (Information-Seeking Strategies)三阶段的深化历程。Moves 关注较为零散的信息搜寻动作;Tactics 对搜寻方法进行了系统分类,也被称为信息搜寻行为层次;ISSs 对搜寻策略进行多维建构。Moves 是低层次的个体行动,涉及搜寻动作过多,研究难度大;而 ISSs 建立了一个由用户、资源和方法组成的三维结构,该整合框架中各要素的构成尚有待进一步发现,开展基于 ISSs 层面的研究具有较大的不确定性。因此,目前研究多着眼于具体的 Tactics,既划分了 Tactics 层面搜寻策略的构成,也对具体的搜寻策略展开研究。M. J. Bates[6]率先概括了4种常用的搜寻方法,即检索、文档结构、询问和监控;H. Xie[7]全面归纳了浏览、搜索、跟踪、挑选、咨询、试错6种方法。袁红[8]比较了传统网络环境和社会化媒体环境的信息浏览效率;田梅和朱学芳[9]运用关键事件技术和扎根理论探讨了偶遇策略,并提出了大学生信息偶遇能力培养模型。

搜寻策略的选择及其形成机制也引起了关注。Z. Yue 等[10]利用"行为对"和隐马尔科夫模型对搜寻策略建模,发现协作式和个人搜寻采用不同的搜寻策略;S. Joo 等[11]采用关联分析和多元回归研究了数字图书馆探索式搜寻中的搜寻策略选择。现有研究还初步探索了搜寻策略选择偏好的形成机制。A. W. Lazonder 等[12]发现,搜寻策略受信息组织方式的影响,用户在寻找网站时使用搜索策略多于浏览策略,而在搜寻网站信息时较多使用浏览策略;王知津等[13]指出消费者类型、购买商品类型以及平台信息环境决定了具体检索策略的选择;王广新等[14]分析了学生选择网络信息搜寻策略的差异,认为个体因素、搜索工具与技术、任务情境与类型影响了搜寻策略选择。

2.3 搜寻意图与策略选择关系的研究

对于用户搜寻意图和策略选择之间规律的揭示,国内外研究刚刚起步,研究成果较少。H. Xie[7]根据40位图书馆馆员对搜寻过程的描述,通过文本分析的方法,抽象出意图、实体、方法和资源之间的关系,以此为基础建立了交互式用户信息意图与信息搜寻策略之间的关联模型。袁红等[15]以采莓模型为基础,采用基于观察实验的日记研究法,通过文本分析和数据建模,揭示了随性休闲搜索行为中查询变化(意图)和思考(策略)之间的规律。但显然,样本选择以及典型案例研究具有一定局限性,使得上述研究成果的科学性有待进一步验证。

综上,尽管信息搜寻意图和搜寻策略领域的研究已相对成熟,但仍存在以下不足:①搜寻意图和搜寻策略分类体系众多,但与具体研究问题相结合时,多维分类体系的完备性和各类目之间的独立性尚不确定;②现有研究证实了搜索策略选择偏好的存在并初步分析了成因,但用户认知与具体实施规划之间的关联机制揭示不够深入,也缺乏真正从用户中心出发,从不同搜索意图角度探讨用户搜索策略运用的研究。可以认为,搜寻意图和搜寻策略分类体系的研究成果仍会在不断涌现的研究问题中得以继承与发展,同时,搜寻意图和搜寻策略选择之间的关联机制是一个亟待发掘的研究领域,以上研究成果为该领域研究问题的凝练和研究的开展提供了重要的理论和方法支撑。

3 基本理论范畴界定

3.1 信息搜寻意图划分

以 A. Broder 理论[1]为框架,根据 D. E. Rose[16]对搜寻意图分类体系的细化思路,本研究将信息类搜寻意图细化为信息查找意图、知识获取意图和建议咨询意图,其中信息查找意图主要体现为寻找动作,包括查询、查找、搜索、列出等,知识获取意图表示想了解某些专门领域的知识,建议咨询意图表示想获得某些意见。本研究同时将事务类搜寻意图具体描述为资源下载意图,将导航类搜寻意图具体描述为导航/URL 意图,见表1。

3.2 信息搜寻策略划分

M. J. Bates[6]提出,信息搜寻的策略主要有搜索(searching)、浏览(browsing)、追踪(monitoring)和偶遇(encountering)。随着社会化搜寻的飞速发展,在问答社区和各类互动平台上向他人寻求帮助逐渐成为一种新的信息获取途径,这就是询问策略。基于此,本研究将信息搜寻策略分为5类:①搜索。用户将自身信息需求提炼为关键词,搜索引擎将搜索结果以一定的排序方式呈现给用户。②浏览。当用户处于一定的信息环境时,会根据自己的信息需求逐一查看信息,并从中分辨出对自己有用的信息。③追踪。对感兴趣的信息定期关注并准确加以捕捉获取,得到与已有的信息需

表 1　信息搜寻意图界定

意图分类	D. E. Rose 等的理论成果[16 – 18]	本文界定
信息类	有指导性(directed)):用户想知道关于某个话题的特定信息	信息查找意图
	无指导性(undirected):用户想了解关于某个话题的任何信息	
	列表(list):用户想得到一组可信的站点或者页面的列表	
	位置(locate):用户想查找某产品或服务的具体地理位置	
	知识(knowledge):用户想获取某些专门的知识	知识获取意图
	建议(advice):用户想得到一些建议、指南或其他方面的指导	建议咨询意图
事务类	下载(download):用户想把某个资源下载到本地或者其他的设备上	资源下载意图
	结果页(result page)):用户想从搜索引擎结果页中打印、保存或阅读资源	
导航类	用户想直接到达一个特定网站	导航/URL 意图

求相关的信息,如微信中的订阅号。④偶遇。计划之外或意料之外的信息获取。J. Krikelas[19] 提出的“偶然信息收集(casual information gathering)”被认为是偶遇概念的起源。⑤询问。使用社会化途径咨询他人意见建议,向具有领域知识的专家或同样经历的人寻求帮助。

4　研究设计与实施

4.1　研究方法

本研究主要采用问卷调查法。信息行为研究一向重视利用科学的程序解决问题,以应对行为主体可能给研究本身带来的主观性、复杂性、不确定性甚至是不真实性。在众多的社会科学研究方法中,问卷一直备受青睐,是目前信息行为研究中使用最为广泛的方法之一[20]。

当然,搜索日志分析、行为实验等方法越来越多应用于信息行为研究。其中,搜索日志分析可以系统分析自然形成的行为痕迹数据,行为实验可以有效创设模拟情境,有效控制研究变量,但它们也有着天然的缺陷,搜索日志不易获取且分析处理技术难度大,行为实验的复杂性使其研究样本规模受限,而问卷调查具有很好的可操作性,易于实现大样本研究。

4.2　问卷设计

4.2.1　搜寻意图与策略选择的题项设计　由于搜索、浏览、询问、追踪一般有明确的信息需求,本研究以问题的形式设定不同的搜寻意图。问题选项 A 到 D 分别对应着某一特定搜寻意图下的上述 4 种搜寻策略,从而获取除偶遇外的 4 种搜寻意图的搜寻策略选择偏好数据。

(1)问题设计。问题设计遵循以下原则:①避免使用专业术语描述搜寻意图。由于调查对象不一定知晓信息搜寻的专业理论和概念,有必要在问题中创设不同的信息需求情境来阐释相应的搜寻意图。如描述信息查找意图时,创设了“如果您想了解有关春节的相关习俗”的需求情境。②尽量全面反映同一意图的不同搜寻情境。每种意图设计 2 种情境,分别为工作学习情境和生活情境,如建议咨询意图的工作学习情境为“假如您对艺术学院的专业感兴趣,想了解是否能转专业以及转专业的流程”,生活情境为“假如您因工作需要去参加一个重要的宴会,您想了解这种场合下黑色西装应该搭配什么样式的衬衫和领带”。③不同意图下的每种情境分别设计单项选择和多项选择。单项选择用于调查单一搜寻策略的选择频次,多项选择用于探索搜寻策略选择的组合。单项选择的表述是“若进行一次查询,您认为采用哪种方法会最终解决这一问题?”,多项选择的表述是“为了得到满意的结果,您倾向于尝试以下哪些方法?”

(2)选项设计。选项设计遵循以下原则:①描述搜寻策略时,关键词有两种:一种从搜寻意图的描述中直接提取,另一种是字面描述的关联词,以兼顾被调查者不同的信息需求和需求表达能力。例如,在了解春节相关习俗一题中,搜索关键词有直接提取的“春节习俗”,还有与之相关的“红包”“拜年”等。②避免被调查者将浏览策略误解为“阅读”,浏览策略描述举例:在新华网、腾讯新闻网等相关网站按照地区、事件类型等分类目录找到相关事件。③询问是一种经由渠道进行的信息传递行为,选项描述时尽可能多地包含各类平台,如百度知道、知乎、Quora、Tipask & Question 等国内外知名问答社区、相关论坛、贴吧等,此外,在涉及专业性比较强的搜索任务时,在选项中加入了“直接向专业人士咨询”。④在询问选项描述时,突出了信息推送平台的多样性,包括订阅号、兴趣部落、社区等推送信息以及电子邮件推送信息。

4.2.2　不同意图下信息偶遇研究的问题设计　偶遇是指用户偶然发现需要或感兴趣的信息,其不确定性决定了只能通过主观题获取相关研究数据。首先简要介绍信息偶遇现象,使被调查者对这一概念有直观的了解;然后适当举出信息偶遇的例子,方便被调查者判断过去的信息获取经历中哪些属于信息偶遇;最后被调查者通过文字描述发生信息偶遇的过程,主要包括当时想要解决的信息问题是什么,即用户的搜寻动机

(用于研究哪些类型的搜寻意图容易发生偶遇现象),以及偶遇的是什么信息(用于判断用户此次信息搜寻经历是否为信息偶遇)。

4.3 调查实施

通过 36 份试调查及时修正问卷中的排版、内容设计及语言表述问题后,开始正式调查。正式调查在问卷星上进行,调查周期为 20 天,回收问卷 245 份,有效问卷 204 份,有效率为 81.9%。研究对象为来自不同地区、各个年龄阶段以及不同教育背景的网络用户。

问卷信效度检验结果显示,α 系数为 0.735,表明问卷数据的一致性较好;KMO 样本测度值为 0.777,Bartlett 球体检验的 2x 统计值为 319.345,自由度为 45,显著性系数为 0,表明该量表适合进行因子分析。以上结果表明本研究问卷设计合理,体现了问卷调查获取的行为数据的准确性、可靠性。

4.4 数据分析工具

为了直观展示不同意图下用户搜寻策略的选择,利用可视化分析工具 UCINET 绘制关系图以确定用户的策略选择偏好;由于传统的数理统计分析无法发现表面数据背后隐藏的规律,利用数据挖掘软件 SPSS modeler 的 Apriori 算法深入探寻不同搜寻意图下用户信息搜寻策略的组合方式。Apriori 算法是一种挖掘关联规则的频繁项集算法,它使用频繁项集的先验知识,使用逐层搜索的迭代方法,找出满足最小支持度的所有频繁项集,再从所有频繁项集中生成关联规则,保留满足最小置信度的规则,即找出所有强关联规则[21],从而从大数据集中发掘关系;现有研究中,关联机制的揭示大多是通过文本分析、内容分析实现,因此利用文本挖掘软件 ROST CM 对不同搜寻意图下的信息偶遇行为的文本数据进行词频统计,探寻信息偶遇的"高发"搜寻情境。

5 数据分析

5.1 不同搜寻意图下用户搜寻策略选择偏好分析

调查问卷中的单项选择题用于确定不同搜寻意图下用户信息搜寻策略的选择偏好,图 1 运用 UCINET 展示了 4 种搜索策略的选择情况。中间位置的方框表示 4 种搜寻策略,边缘位置的方框表示 5 种搜寻意图;方框之间的连线代表选择关系,连线上的数值是选择频次,频次越高,中间位置的策略方框就越大。图 1 显示,搜索和浏览是用户信息搜寻时最常使用的两种搜寻策略。

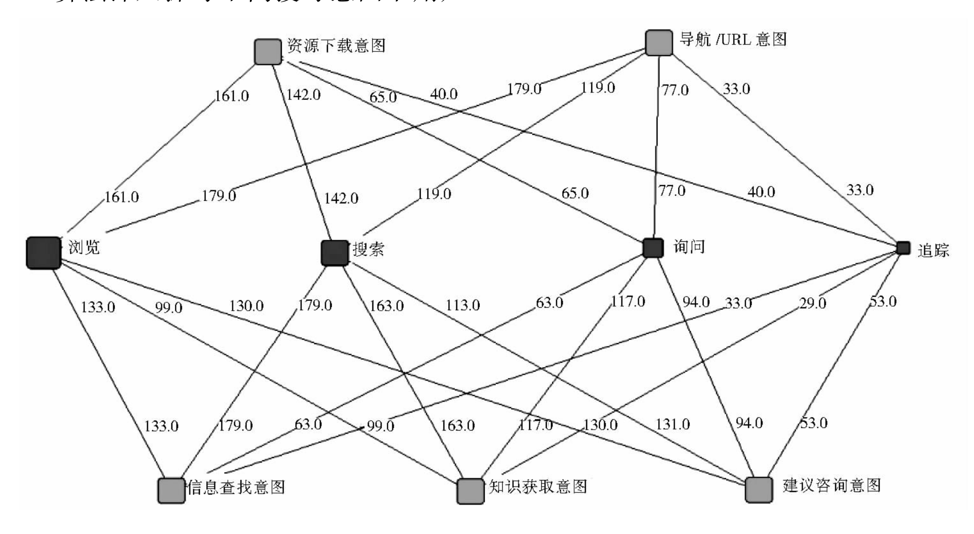


图 1 信息搜寻意图与信息搜寻策略选择关系

统计不同搜寻意图下使用频次最多的搜寻策略,结果见表 2。

从表 2 可知,对于信息查找、知识获取意图,搜索是用户最主要的策略选择,但是对于具有明确个人习惯偏好以及专业性较强的信息问题,搜索策略的有效性不够突出,如对于信息分类组织十分明确的新闻阅读而言,浏览的方式更易于接近个人感兴趣的信息,对于法律知识这类专业性强的问题,询问专业人士则是用户普遍选择的搜寻策略;对于建议咨询、资源下载和导航/URL 意图,浏览是最主要的策略选择,尤其是导航/URL 意图,用户通过特定的网址导航目录的分类就能比较容易快速地得到相应的网址或者超链接。但对于需求表达得特别明晰、易于提炼关键词的建议咨询、资源下载意图(如找电影),搜索策略使用更多。

5.2 不同搜寻意图下用户搜寻策略选择组合分析

调查问卷中的多项选择用于探究不同搜寻意图的

表 2　不同搜寻意图对应的常用搜寻策略

搜寻意图	查询主题	策略选择	频次	百分比(%)
信息查找意图	春节习俗	搜索	110	54.2
	新闻阅读	浏览	88	43.3
知识获取意图	感冒成因	搜索	98	48.3
	房产交易	询问	76	37.4
建议咨询意图	转专业	浏览	73	36.0
	西装搭配	搜索	67	33.0
资源下载意图	备考资料	浏览	106	52.2
	科幻电影	搜索	91	44.8
导航/URL 意图	会计证书考试网站	浏览	83	40.9
	美国电子政务网站	浏览	96	47.3

用户会同时使用哪些搜寻策略。为了更为直观地呈现搜寻策略之间的关联关系,首先利用 SPSS modeler“网络图”进行定性关联分析,生成关联模型如图 2 所示。图中节点代表 4 种搜寻策略,节点之间的连线代表不同策略之间的关联关系,连线的粗细代表相关性的强弱。以图 2 中的建议咨询意图为例,网络图摘要显示,搜索与浏览、搜索与询问之间链接数分别为 253、252,因此它们之间的连线最粗,表明这三种策略具有更强的相关性,浏览与询问之间的链接数为 187、因此这两个节点之间连线比较粗,浏览与追踪、搜索与追踪之间链接数分别为 159、146,因此它们之间的连线较细,而询问和追踪之间的链接数最少,为 138 次,因此这两个节点之间的连线最细,几乎不可见,说明在建议咨询意图下,询问和追踪策略之间的关系性较弱。

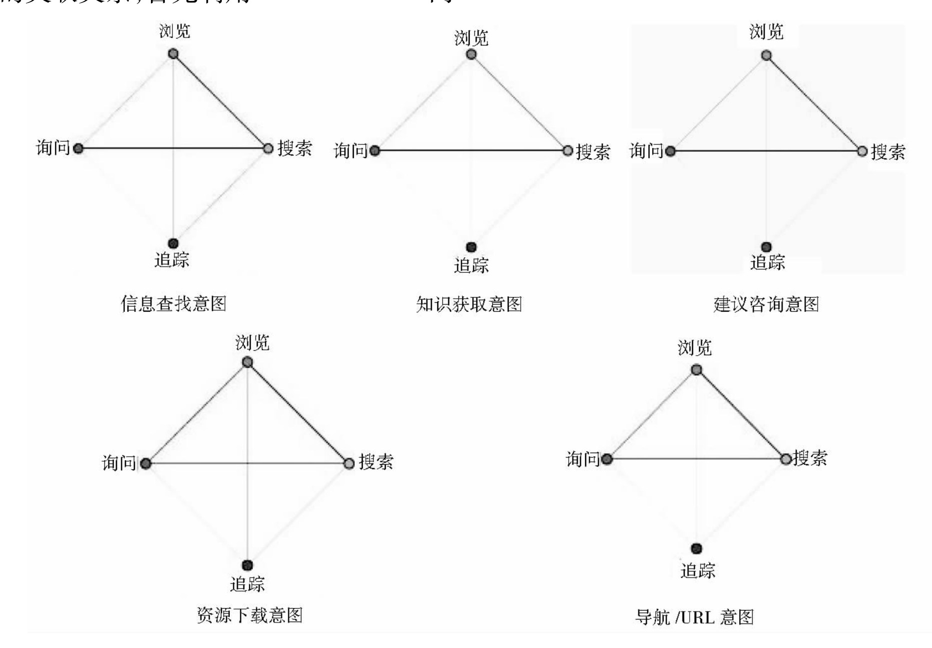


图 2　搜寻意图与搜寻策略之间的网络图关联模型

网络图形式的关联模型尽管直观,但没有产生严格的关联规则。SPSS Modeler 提供了专业的深入探索和挖掘数据模型的算法。本研究运用 SPSS Modeler 中的 Apriori 算法进行关联分析,具体流程为:①导入数据文件,根据要求设置字段类型,Apriori 算法只能对布尔型数据进行分析,因此把所有字段类型设置为“标志”,角色设置为“任意”。②建立 Apriori 算法模型,支持度设置为 10%,置信度设置为 80%,支持度(support)指事件 A 和事件 B 同时发生的概率,用以衡量规则需要满足的最小重要性;置信度(confidence)指发生事件 A 的基础上发生事件 B 的概率,它表示关联规则需要满足的最低可靠性。同时将最大前项数设置为 3。③运行并查看 Apriori 关联模型结果。图 3 为信息查找意图下搜寻策略组合的数据挖掘结果。



后项	前项	支持度百分比	置信度百分比
搜索	浏览	66.502	81.481
搜索	询问	59.852	91.358
搜索	询问 浏览	33.99	88.406
搜索	追踪 询问	21.429	93.103
搜索	追踪 询问 浏览	14.039	98.246

图 3　信息查找意图下策略组合数据挖掘结果

再依次进行知识获取、建议咨询、资源下载和导

航/URL 意图下搜寻策略选择的关联分析,从产生的所有关联规则中筛选支持度大于 50% 的规则。支持度阈值设定过低,将产生大量关联规则,适当提高支持度,有利于挖掘出重要的强关联规则。Apriori 算法运行产生的全部支持度大于 50%、置信度大于 80% 的 9 条关联规则见表 3。

表 3 Apriori 算法运行结果

搜寻意图	后项	前项	支持度(%)	置信度(%)
信息查找意图	搜索	浏览	66.502	81.481
	搜索	询问	59.852	91.358
知识获取意图	搜索	询问	67.98	81.522
建议咨询意图	搜索	浏览	71.921	86.644
	搜索	询问	70.197	88.421
资源下载意图	浏览	追踪	55.419	82.222
导航/URL 意图	搜索	浏览	78.818	82.5
	搜索	询问	72.906	85.135
	搜索	询问 and 浏览	55.665	80.973

表 3 展示了不同意图下用户搜寻策略选择的大概率组合。在信息查找意图中,{搜索、浏览}、{搜索、询问}的支持度分别是 66.502% 和 59.852%,置信度分别为 81.481% 和 91.358%,说明用户在选择搜索策略的同时存在比较大的可能选择浏览和询问策略;同样,知识获取意图频繁项集为{搜索、询问},建议咨询意图的频繁项集为{搜索、浏览}和{搜索、询问};资源下载意图的频繁项集为{浏览、追踪};导航/URL 意图的频繁项集为{搜索、浏览}、{搜索、询问}、{搜索、询问、浏览}。

5.3 不同搜寻意图下信息偶遇分析

将被调查者提交的自身经历的信息偶遇现象的描述文本运用 ROST CM 进行字段抽取,并按行进行一般性行处理,提取规范文本中描述搜寻动机的文本,并对其进行词频分析,将词频统计结果用可视化方式展示,如图 4 所示。图中字体越大表明该词汇出现的频率越高。

图 4 信息偶遇时信息搜索动机词频分析

将以上搜索动机归入本研究界定的 5 类搜寻意图并统计其频次,降序排列结果如表 4 所示:

表 4 搜寻意图与搜索动机对应关系

搜寻意图	搜寻动机	百分比(%)
信息查找意图	寻找做菜信息、查找户外服装、查找科比退役视频、查找歌曲、查找单词、查找考研面试信息、查找减肥方法、查找资料、浏览网页、浏览新闻、查找租房信息、查找课程学习相关信息、查找商品信息、查找文献	68.3
资源下载意图	查找下载课题资料、查找下载 Python 数据挖掘包、查找下载电影、查找下载图片	15.0
建议咨询意图	浏览微博中的用户推荐、在问答社区向他人求助	8.3
导航/URL 意图	获取网店网址	5.0
知识获取意图	了解碘含量高的食品有哪些、了解网站建设的知识	3.4

表 4 说明,具有明确需求的信息搜寻时,往往会获取到无预期或低预期的信息,用以解决已有的问题或满足兴趣[22-23],这种信息偶遇现象绝大多数发生在用户出于信息查找意图的信息搜寻过程中,其次是资源下载意图的信息搜寻也会偶遇有价值的但并非当时情境所期待的信息。此外,在社会化媒体平台(如问答社区、论坛、贴吧等)进行建议咨询意图的信息查找时也时常会信息偶遇。

6 结果讨论

6.1 无论搜寻意图的类型如何,搜索和浏览策略均最多被选择

CNNIC 研究报告显示,截至 2019 年 6 月,搜索引擎用户规模达 6.95 亿,使用率为 81.3%[24]。某种程度上,搜索已经成为网络信息搜寻的代名词。一方面是源自于搜索引擎可以有效匹配用户的搜索请求,帮助用户形成了“有事搜一搜”的行为模式。随着搜索引擎技术向移动化、个体化、实时化、情境化发展,以及地理感知搜索、多媒体搜索、跨语言搜索等新技术的广泛应用,搜索引擎在识别用户搜寻意图、进而帮助用户解答问题或形成对特定问题的理解上表现更为出色,更是使搜索成为用户和信息之间的“最佳实践”。最多选择搜索策略的原因还在于用户在漫长的搜寻探索过程中常见的查询重构行为,也会带来搜寻策略选择频次的增加。

浏览策略常见两种方式,一是分类浏览,它以分类

目录为中介,浏览的信息有明确的分类主题;二是关联浏览,它以相关资源为中介,浏览的信息具有相关主题或围绕共同兴趣。无论哪种方式,它都使用户在网络空间中从一个资源向另一个资源游弋,过滤不必要的信息,从而实现其信息需求。浏览策略在查准和启发知识创新方面独具优势,已成为与搜索同等重要的策略。

6.2 信息查找、知识获取意图偏好搜索策略,建议咨询、资源下载、导航/URL 意图偏好浏览策略

尽管搜索和浏览为用户信息搜寻的基本策略,但不同类型搜寻意图的策略选择偏好仍存在差异。从网络图中搜索和浏览节点具有的链接数的比较可以看出,对于信息查找、知识获取、建议咨询意图,搜索最有效、最多被选择;而对于资源下载、导航/URL 意图,浏览是最主要的搜寻策略。这主要是因为搜索和浏览策略的差异性所致。

搜索策略的运用需要消耗人的认知资源从记忆中唤起搜索词来表达信息需求,并借由搜索引擎工具加以实现,而浏览策略具有信息遍历的特征,用户会多次浏览"中心"网站,并探索"相关"网站,多次交互、多次循环、变化目标,被认为是曲折和费时的。因此,搜索策略的运用能更好地满足信息查找、知识获取对于搜寻效率的要求,帮助用户以尽可能小的努力获取信息。而浏览能通过超链接、相关资源的中介作用以及用户间的相互作用(评价、推荐、关注、讨论等)使相对模糊的建议咨询、资源下载、导航/URL 需求明确化,帮助用户高效地找到感兴趣且适合的信息。

6.3 无论搜寻意图的类型如何,{搜索、浏览}、{搜索、询问}策略组合最多被选择

信息搜寻是一个受多因素影响的复杂系统,不同的"进入"策略以及后续的策略组合会产生不同的信息搜寻效果。同时,海量信息和冗余信息带来的搜寻难题更需要运用综合的搜寻策略来应对。实际上,搜索和浏览不能解决全部的信息问题,比如,不可能通过搜索或浏览搜寻到个人或组织的所有知识。因此,要达成"最大满意",必须付出更多的主动搜寻努力。搜索、浏览、询问均为主动搜寻策略,{搜索、浏览}、{搜索、询问}的常见策略组合说明了用户在信息搜寻活动中的主导地位,他们倾向于选择主动搜寻,而不是被动等待信息的靠近。

{搜索、浏览}的策略组合最为常见,信息搜寻通常是"搜索→浏览(结果)"的循环过程。首先,无论何种搜寻意图,搜索引擎搜索、构造检索词、浏览页面、点击页面链接、浏览新的页面都是用户启动搜寻的常规模式,搜寻进程中,如果搜索引擎返回的结果不令人满意,用户将修改其查询词以优化搜索结果,再继续"搜索→浏览(结果)"的行为序列。由该组合模式代表的从快速密集搜索开始,到中后阶段的细致搜索和浏览,是最为典型的搜寻行为过程。{搜索、询问}策略组合产生于社会化媒体的关系和互动,带有最为显著的社会化的烙印。询问策略从本质上更易于针对特定的信息需求集思广益,群策群力,并借由社会化这一有效的信息传递及筛选机制提升搜寻效果。

6.4 知识的特征决定了知识获取意图搜寻中{搜索、询问}是最显著的策略组合模式

知识获取意图是深层次信息类意图,它不仅是扫描、定位和发现知识,更是一系列的学习探索和交互,其实现路径往往是搜索、浏览、询问、追踪策略的综合运用。在搜寻初期,用户搜索频率低,浏览网页数目多,以求发现完备的主题知识,之后会频繁更换搜寻策略、细致浏览搜索网页以形成系统的知识结构,随着搜寻进程的深入,用户可能会选择询问来评价、交流对知识的理解,也会通过追踪保持知识储备以应对不断出现的信息问题。

除了知识获取意图具有策略选择多样化的特点,{搜索、询问}策略组合更是体现了知识获取信息搜寻中人机交互策略和人际交互策略的融合。当知识以图形、文字、公式等编码形式存在时,搜索策略操作简单、响应快速,用户能通过最少的搜索、最简单的关键词得到预期的搜索目标。然而,当知识以个体经验、感觉、意识、判断、灵感、创意等形式存在时,搜索或浏览策略对这种难以表达和传递的信息获取就显得力不从心了,代表着人际交互的询问策略更能发挥所长,开展更有针对性的知识搜寻。

6.5 {浏览、追踪}是资源下载意图信息搜寻的基本策略组合单元

资源下载意图一般包括两个层次,其一,获取某一特定资源,其二,对已确定目标资源的接收和利用。资源下载意图最接近信息的使用,多产生于搜寻后期的评价和选择阶段,此时用户的不确定性减少,搜索引擎的全网搜索快速匹配优势难以体现,反而是浏览策略可以过滤掉不必要的信息,能更有效地作用于资源的选择,再通过信息使用调整预先设定的目标,直至获得目标资源。

{浏览、追踪}策略组合在资源下载意图信息搜寻中经常出现的原因在于,资源下载意图往往难以通过

一次搜寻有效满足,表现在获取的资源在数量、质量、广度、深度上的不足,任务情境和资源本身的动态发展变化同时在加深这种“不足之惑”和“求足之愿”,因此用户会长期保持一种针对该资源需求的倾向性和敏感性,当相关信息出现时就会加以捕捉,这就是资源下载中的追踪策略。

6.6 {搜索、浏览、询问}策略组合发生在导航/URL意图所处的多种搜寻情境下

导航/URL 意图是搜寻特定网站或包含足够导航信息的网站页面的直接意图,用户在执行检索时已在头脑中形成了查询意向,知道或者认为存在网址可以满足自己的信息需求,因此这类搜寻意图相对简单,多种策略都可以有效搜寻。而{搜索、浏览、询问}策略组合的运用与用户搜寻意图中导航/URL 意图的特殊性有关。

一方面,信息查找意图会因信息甄别或信息聚集的需要而转化为导航/URL 意图,如苹果 XR→苹果官网→苹果 XR 官方介绍的搜索路径体现了信息查找意图→导航/URL 意图→知识获取意图的变化。同时,资源下载、知识获取等意图的搜寻通常也需要先导航到某个确定的网站,如用户在搜索框输入驱动之家→驱动之家网站查找特定驱动→下载驱动的搜索活动体现了导航/URL 意图→信息查找意图→资源下载意图的变化,可以认为,导航/URL 意图往往是其他意图的源头或桥梁。{搜索、浏览、询问}策略组合就发生在导航/URL 意图所处的多种搜寻情境下,如浏览推荐内容后的进一步搜索;或是用户经过长时间搜索且无法准确判断网络信息定位时会引起认知倦怠,通过询问来协助评价,同时快速结束搜索;或是关键词搜索后逐一浏览搜索结果集。

6.7 信息查找、资源下载和建议咨询意图易于触发信息偶遇

偶遇是缺乏目的性的被动搜寻策略,同时也是一种重要的信息获取方式。M. J. Bates[6]曾估计人们获得的信息中超过 80% 是在日常活动中意外发现的。已有研究表明,信息偶遇受个人差异和环境特征的影响,但偶遇的发生不以信息需求为前提[23]。本研究认为,某些特定类型的搜寻意图更易于触发信息偶遇,显然与此相矛盾。可能的解释是:这与本研究对搜寻意图的划分有关,信息查找、资源下载和建议咨询意图代表了最多的搜寻情境,发生偶遇的概率自然增加。还有的合理解释是偶遇的发生与用户操作行为的频繁程度相关,如信息查找意图会频繁提交查询,会话次数较多,偶遇的可能较大,而导航/URL 意图下用户倾向于进行为数不多的点击,偶遇的可能性较小。此外,被调查者描述的建议咨询意图下偶遇发生的情境都与社会化互动相关,这说明社会化媒体有利于社会导航和信息探索,能够有效地支持信息偶遇。

7 结语

本文通过问卷获取用户信息搜寻意图与策略选择行为数据,通过数据挖掘和文本分析建立信息搜寻意图与信息搜寻策略之间的关联机制,这为优化搜索系统的功能设计提供了一些思路。如搜索系统应首先强化其对搜索和浏览功能的支持,在此基础上,深入了解用户在不同意图下的策略组合关联关系,并据此改进功能区的布局,优化和改进支持探索、支持社会化交互,帮助用户建立有效信息关联的界面和功能设计,从而提升用户的搜寻体验和绩效。

本文的不足之处在于,问卷调查的研究方法难以避免被调查者对问题认识的主观性和表述偏差,其调查结果虽全面但难以深入,后续研究将探索用户搜寻行为数据的自动化获取方法,结合问卷调查和日志分析,扩大分析数据集。此外,本文仅对 Tactics 层面的搜寻策略展开研究,而从 Moves 和 ISSs 层面全面探究用户搜寻意图与策略之间的关联机制也将在后续研究中开展。

参考文献:

[1] BRODER A. A taxonomy of Web search[J]. ACM SIGIR forum, 2002, 36(2): 3-10.

[2] NGUYEN V B, KAN M Y. Functional faceted Web query analysis [C]//WWW 2007. Proceedings of the 16th international conference on World Wide Web. New York: ACM, 2007: 32-39.

[3] BELKIN N J, KELLY D, KIM G. Query length in interactive information Retrieval[C]//Proceedings of the 26th annual international ACM conference on research and development in information retrieval. New York: ACM, 2003: 205-212.

[4] VALLET D, ZARAGOZA H. Inferring the most important types of a query: a semantic approach[C]// Proceedings of the 31st annual international ACM SIGIR conference on research and development in information retrieval. New York: ACM, 2008: 857-858.

[5] 张晓娟. 信息类、导航类与事务类查询的网络动态性分析[J]. 现代图书情报技术, 2017, 1(4): 13-23.

[6] BATES M J. Toward an integrated model of information seeking and searching[J]. New review of information behavior research, 2002(3): 1-15.

[7] XIE H. Patterns between interactive intentions and information-seeking strategies[J]. Information processing & management, 2002,

38(1): 55 – 77.
[8] 袁红. 社会化媒体环境下用户信息浏览策略的效率评价研究[J]. 情报理论与实践, 2017, 40(4): 89 – 94.
[9] 田梅, 朱学芳. 基于现行信息素养模型的大学生信息偶遇能力培养[J]. 图书情报工作, 2015, 59(17): 41 – 46.
[10] YUE Z, HAN S, JIANG J, et al. Search tactics as means of examining search processes in collaborative exploratory Web search [C]// Proceedings of the 5th Ph. D. workshop on information and knowledge ACM. New York: ACM, 2012: 59 – 66.
[11] JOO S, XIE I. How do users' search tactic selections influence search outputs in exploratory search tasks? [C]// Proceedings of the 13th ACM/IEEE-CS joint conference on digital libraries. New York: ACM, 2013: 397 – 398.
[12] LAZONDER A W, BIEMANS H J A, WOPEREIS I G J H. Differences between novice and experienced users in searching information on the World Wide Web [J]. Journal of the American Society for Information Science, 2000,51(6): 576 – 581.
[13] 王知津, 韩正彪, 周鹏. 电子商务网站顾客信息搜寻行为形成机制研究[J]. 图书与情报, 2011(3): 12 – 16.
[14] 王广新, 谭金波. 学生网络信息搜索策略选择的差异分析[J]. 情报资料工作, 2014(5): 56 – 61.
[15] 袁红, 王奥. 采莓与搜索:随性休闲搜索行为特征研究[J]. 现代情报, 2019(4): 59 – 68.
[16] ROSE D E, LEVINSON D. Understanding user goals in Web search[C]// WWW 2004: Proceedings of the 13th international conference on World Wide Web. New York: ACM, 2004: 13 – 19.
[17] 齐富民,谢晓尧,吴静. 网络搜索中用户搜寻意图识别的研究[J]. 计算机工程与设计, 2014, 35(4): 1285 – 1292.
[18] 唐静笑, 吕学强, 柳成洋,等. 用户查询意图的层次化识别方法[J]. 现代图书情报技术, 2014(1): 36 – 42.
[19] KRIKEIAS J. Information-seeking behavior: patterns and concepts [J]. Drexel library quarterly, 1983, 19(2): 5 – 20.
[20] 姜婷婷, 杨佳琪, 李倩. 信息行为领域概念空间构建与研究进展述评[J]. 图书情报知识, 2019(1): 99 – 108.
[21] 张红艳, 都娟. 关联规则中 Apriori 算法的应用[J]. 数字技术与应用, 2011(8): 14 – 15.
[22] ERDELEZ S. Information encountering: it's more than just bumping into information[J]. Bulletin of the Association for Information Science and Technology, 1999, 25(3): 26 – 29.
[23] MCCAY-PEET L, TOMS E G. Investigating serendipity: how it unfolds and what may influence it[J]. Journal of the Association for Information Science & Technology, 2015, 66(7): 1463 – 1476.
[24] 中国互联网络信息中心. 第 44 次中国互联网络发展状况统计报告[R/OL]. [2019 – 08 – 30]. http://cnnic.net/hlwfzyj/hlwxzbg/hlwtjbg/201908/P020190830356787490958.pdf.

Research on the Correlation Mechanism Between Users' Seeking Intention and Seeking Strategy Selection

Yuan Hong
School of Economics and Management, Nantong University, Nantong 226019

Abstract: [**Purpose/significance**] Seeking intention is the starting point of search process, seeking strategy is the planning of search behavior, and the correlation mechanism between seeking intention and seeking strategy selection is the reorientation of the dynamic development of search behavior. It is of great significance of exploring the correlation mechanism between information seeking intention and seeking strategy selection to deeply understand the causes and processes of the changing search behavior. [**Method/process**] This paper collected the data of seeking intention and seeking strategy through questionnaires, used UCINET to study the preference of strategy choice, used SPSS modeler to explore the most commonly used strategy combinations, and used ROST CM to study information encounter under different intentions. [**Result/conclusion**] Searching and browsing are the two most frequently used seeking strategies. Users have obvious differences in their preferences for different seeking strategies and different seeking strategy combinations corresponding to different seeking intentions. For the strategy of information encountering, it is more likely to occur in the situation of information searching, resource downloading, and advising and consulting.

Keywords: seeking intention　seeking strategy　correlation mechanism